AF TECHNICAL REPORT No. 5914, Part 3 May 1950

DESIGN OF INSTRUMENT DIALS for MAXIMUM LEGIBILITY

Part 3. Some Data on the Difficulty of Quantitative Reading in Different Parts of a Dial

William E. Kappauf
and
William A. Smith
M.

Princeton University

United States Air Force
Air Materiel Command
Wright-Patterson Air Force Base, Dayton, Ohio

NOTE

When drawings, specifications, and other data prepared by the War Department are furnished to manufacturers and others for use in the manufacture or purchase of supplies, or for any other purpose, the Government assumes no responsibility nor obligation whatever; and the furnishing of said data by the War Department is not to be regarded by implication or otherwise, or in any manner licensing the holder, or conveying any rights or permission to manufacture, use, or sell any patented inventions that may in any way be related thereto.

The information furnished herewith is made available for study upon the understanding that the Government's proprietary interests in and relating thereto shall not be impaired. It is desired that the Patent & Royalties Section, Office of the Judge Advocate, Air Materiel Command, Wright-Patterson AFB, Dayton, Ohio, be promptly notified of any apparent conflict between the Government's proprietary interests and those of others.

Espionage Act

Notice: "This document contains information affecting the national defense of the United States within the meaning of the Espionage Act, 50 U. S. C., 31 and 32, as amended. Its transmission or the revelation of its contents in any manner to an unauthorized person is prohibited by law."

"The above Espionage Notice can be disregarded unless this document is plainly marked with a security classification as "Restricted," "Confidential," "Secret," or "Top Secret."

The U. S. Government is absolved from any litigation which may ensue from the contractor's infringing on the foreign patent rights which may be involved.

AF TECHNICAL REPORT NO. 5914, Part 3 May 1950

DESIGN OF INSTRUMENT DIALS for MAXIMUM LEGIBILITY

Part 3. Some Data on the Difficulty of Quantitative Reading in Different Parts of a Dial

William E. Kappauf
and
William A. Smith

Parts 1 and 2 of this report were published as memorandum reports

Princeton University

United States Air Force
Air Materiel Command
Wright-Patterson Air Force Base, Dayton, Ohio

Air Force-WPAFB-(A)-O-23 MAY 50 550

FOREWORD

This report was prepared at Princeton University under USAF Contract No. W33-038 ac-14480 covering work on legibility problems in instrument design. The contract was initiated under the research and development project identified by Expenditure Order No. 694-15, and was administered by the Psychology Branch of the Aero Medical Laboratory, Engineering Division, Air Materiel Command, with Dr. Walter F. Grether acting as Project Engineer.

Two previous reports in this series from Princeton University were published as USAF Air Materiel Command Memorandum Reports TSEAA-694-1L, 20 October 1947, and MCREXD-694-1N, 12 July 1948.

ABSTRACT

This is the report of an analysis to determine whether the likelihood of error in quantitative reading is greater in one part of a dial than in another. The error data which are examined are based on over 45,000 dial readings made in the course of two experiments involving a total of 28 subjects.

No evidence is found which would indicate that local scale reading errors (errors of rounding, interpolation, and the like) vary with dial sector. The frequency of systematic scale reading errors depended on dial sector for dials graduated from 0 to 50 or from 0 to 100 but not for dials with scale ranges of 0 to 200, 0 to 400, or 0 to 600. The 50's and 100's dials were similar in that they were numbered by ten unit steps. The dominant error made in reading them was an error of reporting a scale value too great by ten units. This error was more prevalent on the right dial halves than on the left and was especially frequent in the scale region from 0 to 9.

Thus, while sector has no consistent effect on either local errors or systematic errors for many dials, it may influence the occurrence of specific systematic errors on certain scales. As more is learned regarding scale design factors related to systematic errors, the easier it will be to account for the effect of sector on the ten units error here reported or to identify other dial designs which may produce error patterns varying with sector.

PUBLICATION APPROVAL

For the Commanding General:

Walter A. Carlson
WALTER A. CARLSON
Colonel, USAF (MC)
Chief, Aero Medical Laboratory
Engineering Division

CONTENTS

FIGURES AND TABLES

DESIGN OF INSTRUMENT DIALS FOR MAXIMUM LEGIBILITY. III. SOME DATA ON THE DIFFICULTY OF QUANTITATIVE READING IN DIFFERENT PARTS OF A DIAL.

I. INTRODUCTION

Warrick and Grether (1948) and Grether and Connell (1948) have recently reported data on the check reading of fixed-scale, moving pointer instruments which, taken *in toto*, show that deviations of a pointer from an index or stated value are correctly responded to more frequently when the index is on the left half of the dial near the nine o'clock position than when it is on the right half of the dial near the three o'clock position. This result is presumed to follow from our greater familiarity with, and possibly the greater naturalness of, scales where values proceed in increasing magnitude upward.

It is the purpose of the present paper to discuss some observations on the relative difficulty of different portions or sectors of linearly graduated dials when quantitative scale readings are to be made. Data for this form of reading, paralleling the data for check reading, could have bearing on a panel designer's choice of scale orientation for instruments which will typically be read over restricted numerical regions of the scale.

There is one previous research paper which touches on the present question of the accuracy of quantitative scale reading by display sector. This is a report of Garner (1946) in which it is demonstrated that estimates of target bearing on a radar scope vary in accuracy with bearing, being less in error when the target is in the upper or lower quadrants of the scope than when it is in the right or left quadrants. These findings are interpreted to be a consequence of the relation of the operator's line of sight to the scope and its markings. Garner's assumption appears to be that if the display were viewed perpendicular to the operator's line of sight, as is usually the case in instrument use, quadrants or sectors would be of equal difficulty. As applied to interpolation errors, the sort of error with which Garner was principally concerned, this hypothesis is intuitively reasonable. There remain, however, the matter of assembling data in support of this hypothesis and the task of demonstrating whether or not particular systematic reading errors, e.g. reversal errors, or systematic errors as a class are influenced by sector.

The data analyzed in this paper were drawn from two experiments. The first of these was a short study conducted for the purpose of collecting information on the sector problem. Eight subjects were used. They worked with seven types of dials, each of which was read in 100 different pointer positions. The second experiment, the basis of a forthcoming report on dial size and graduation, was one from which supplementary evidence on the sector question could be obtained. In this study, twenty subjects made readings on fifteen different kinds of dials. Some 45,000 readings were involved in the two experiments.

The scales used in these experiments resembled numbered instrument scales. The general intent in preparing them had been to develop a series of typical scales which would make possible a study of graduation scheme and scale length variables. Reading records obtained with these materials proved to be useful for the present purpose because they contained both interpolation errors and systematic, scale interpretation errors, and because they made it possible to look for sector effects under a variety of conditions, in what were essentially replicated sub-experiments with different dial materials. They thus provided a convenient basis for checking the generality of some earlier preliminary results which had suggested the presence of a sector effect in reading data for dials scaled from 0 to 100 (Kappauf, Smith, and Bray, 1947).

For reference purposes, the present experiments will be identified as Experiment 3 and Experiment 4.

II. APPARATUS AND GENERAL PROCEDURE

The apparatus and testing procedure employed in these studies have been discussed in detail in other reports (Kappauf, Smith, and Bray, 1947; Kappauf and Smith, 1948). The apparatus provides for individual testing. With the presentation of each stimulus card, the subject reads 12 identically designed dials, the central 10 of which are test dials. Prior to each card presentation, the subject views a sample dial which indicates the type of dial which is to be read next. As the subject reads, the experimenter records responses against a list of correct values and takes stop clock time on the series of 10 test dials. About fifty cards can be read conveniently in an hour of testing.

III. EXPERIMENT 3

Subjects

Eight college students served as subjects. All evidenced at least 20/20 binocular visual acuity when tested in the experimental apparatus with the Lebensohn near vision acuity chart (Lebensohn, 1936). The test was given at the 28 inch distance used for dial reading, instead of at the usual clinical near reading distance, so in order to qualify as having 20/20 vision, a subject was required to read correctly at least eight of the eleven 20/40 letters and numbers on the chart.

Stimulus Material

The dial reading material presented to the subjects consisted of high contrast photographic prints on mat paper. As illuminated in the test situation, the white dial markings had a brightness of 3 foot lamberts.

The seven different dial designs which were employed are shown in Figure 1. Four types were 100's dials; graduated every 10 units, every 5 units, every 2 units, and every unit. Three were 200's dials; graduated by tens, by fives, and by twos. All were used in two sizes, 2.8 inches diameter and 1.4 inches diameter. A total of 1,400 test dials were involved, 100 for each type and size of dial.

It should be noted that the pointer settings shown on the dials were all held within 0.3 of some exact unit position on the scale (see Kappauf, Smith, and Bray, 1947). This restriction of settings was introduced in order to eliminate ambiguous readings at or near 0.5 positions and reduced somewhat the number of interpolation errors recorded.

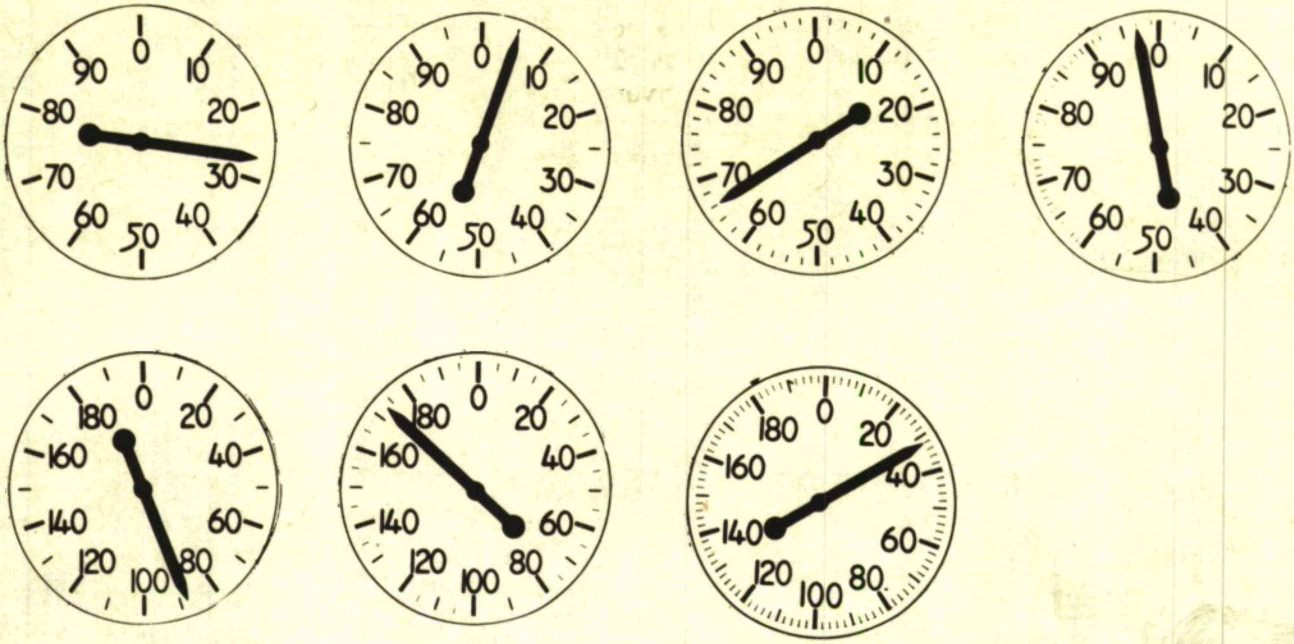

Figure 1. Dial designs used in Experiment 3. Note: These dials are "negatives" of those actually read in the experiment. They are shown here as black on white instead of white on black in order to make reproduction easier.

Plan of the Experiment

In this experiment, pointer settings were distributed systematically around each dial so that the frequency of occurrence of settings involving specific final digits would be exactly balanced within successive dial tenths. For every 100's dial each of the 100 possible unit settings was used exactly once in the test series. For each of the 200's dials, a comparable series of 100 settings was developed using, in each tenth of the dial, a set of 10 distributed readings which involved each final digit once. It was therefore possible to enquire about reading difficulty within dial tenths or within wider regions embracing adjacent dial tenths.

The 80 stimulus cards for the 100's dials (four types of scale, two sizes of dial, and 10 cards for each type-size combination) were divided into two balanced test decks of 40 cards each. For each card in the first deck, a companion card containing the same dials in the other size appeared in the second deck. Within each deck, dial types and sizes were rotated in systematic fashion. Two similar test decks were prepared from the sixty 200's dial cards. Sixteen additional 100's dial cards and 12 additional 200's dial cards were used as practice material.

On a preliminary practice day, each subject was familiarized with the test routine and read all the material in the practice deck. During the four days of the experiment proper, each subject read through the four test decks, one deck each day. The test decks were given in a different balanced order to each subject. Each day the test deck readings were preceded by a brief warm-up run on half the practice cards bearing the kinds of dials to be read that day. Test sessions were approximately one hour in length. They were interrupted with two minute rest periods at the conclusion of the warm-up readings and after each quarter of the test readings. With few exceptions, the five sessions for each subject were run on consecutive days.

The subjects were instructed to read the dials as rapidly as possible and to make each reading to the nearest unit. Accuracy of reading, as such, was not mentioned in the instructions, but emphasis was placed on "reading to the nearest unit." Instructions were reviewed before each test session.

Handling of the Data

The reading errors recorded in the course of the experiment have been classified and summarized in Table I.

TABLE I

GENERAL BREAKDOWN OF READING ERRORS IN EXPERIMENT 3

	Total number of readings in error		Percentage of readings in error	
	100's dials	200's dials	100's dials	200's dials
Total readings made	6400	4800	100.0	100.0
Total incorrect readings	433	1019	6.8	21.2
Errors of number identification	11	18	0.2	0.4
Total incorrect scale readings	422	1001	6.6	20.9
Errors of 1 or 2 units	382	908	6.0	18.9
Errors of 3 or 4 units	1	9	-	0.2
Errors of 5 or more units	43	117	0.7	2.4

In an earlier experiment of this series (Experiment 1: Kappauf, Smith, and Bray, 1947), certain reading errors were identified as arising from incorrect reading of the scale numbers in situations where the numbers had been rendered somewhat ambiguous by the way in which the pointer covered them. When the present experiment was being planned, steps were taken which, it was hoped, would prevent the recurrence of such errors. First, certain of the dial numbers were modified to eliminate number confusions, at least to the satisfaction of the experimenters. On the 100's dial, the 5 at the 50 position was radically changed to eliminate confusion with the 6 or 3. Longer tails were added to the 6 and 9. Second, on the 200's dials, where not all of the 200 scale values were to be read in the experiment, those settings which were expected to cause special confusions in number reading were specifically avoided. Third, the schedule of reading for any subject on any one test day was arranged to include 100 dials exclusively or 200 dials exclusively. This, it was hoped, would so familiarize the subjects with the particular number scale on each day's dials that they would not need to rely on reading the individual dial numbers. In spite of these precautions, however, some 0.3% of all the test readings involved errors which were clearly of this number reading sort. Thus a partially masked 20 was occasionally read as 30; 51 as 31; 61 as 51 or 81; 62 as 52 or 82; 137 as 157; 163 as 103 or 183; 165 as 185.

In order that sector effects would not be confused with this problem of number identification, the foregoing 29 errors were separately classified and, except for inclusion in Table I herewith, were eliminated from further consideration.

It should be remarked, however, that the occurrence of errors of this type during test sessions involving dials with but one number scale means that practiced instrument readers probably rely on reading scale number values more than we might surmise.

For purposes of analysis, the scale reading errors proper were classified as "large or systematic" and "local" on the basis of their size. Errors of five units or more, typically errors of 5, 10, 20, 50, and 100, were classed as systematic. Errors of one and two units were classed as local. Errors of three and four units, in addition to being ambiguous, proved to be so infrequent (see Table I) that it was deemed a reasonable simplification to include them in neither category. This means that the reader should interpret the two categories used in the discussion below as somewhat arbitrarily chosen devices for getting indices of the likelihood of scale misinterpretations on the one hand, and of local reading precision on the other.

It will be noted in Table I that the scale reading errors in the three size categories do not sum to the values listed for "total incorrect scale readings." This is due to the fact that certain readings were obvious composites of systematic and local errors and so were tallied in two of the error categories. Particular errors which were so treated were errors of 8, 9, 11, 12, 18, 19, 21, 22, 98, 99, 101, and 102 units.

IV. EXPERIMENT 4

This experiment can be described briefly, for it was run in a manner closely resembling Experiment 3. An important difference was that the sectors within which the pointer settings were balanced in final digit distribution were never so small as dial tenths.

Subjects

Twenty high school students served as subjects. All met the same acuity test standards discussed for the subjects in Experiment 3.

Stimulus Material

The dial cards were the same in form as those described above. There were 15 dial designs in all. The dials were scaled from 0 to 50, 0 to 100, 0 to 200, 0 to 400, and 0 to 600. They were graduated by ten unit intervals, by five unit steps, by two unit steps, or by units, as shown in the list of dial types in Table III below. All designs were used in both 2.8 inch and 1.4 inch diameter sizes.

For each type and size of the 50's dials, 50 pointer settings were used, one at each of the unit positions. For each of the other dials, a series of 60 readings was used. The latter were systematically split as follows: 30 on the right side of the dial and 30 on the left; six ending with each of the 10 possible final digits; six with each of the 10 possible tens digits; and an equal number within each available 100 units range on the total scale. Within this controlling framework, the particular hundreds, tens and final digits combinations used were obtained by working from random number tables.

Plan of the Experiment

Each subject came to the laboratory for six reading sessions. On the first day he was familiarized with the test routine and read practice cards for all dial types to be used in the experiment. On the subsequent five days he spent one test session each on a 50's deck, a 100's deck, a 200's deck, a 400's deck and a 600's deck. The order in which these different decks of cards were read was different for every subject and was balanced so that each deck was read by four of the subjects on each of the five test days.

The test session schedule, the use of warm-up readings and rest periods, and the reading instructions given to the subjects were as described for Experiment 3.

V. RESULTS

The results of a tally of the records of the two experiments are summarized in Tables II and III. In each table, Section A shows the frequency of local errors of one and two units and Section B shows the frequency of errors of five units or more. The sectors listed for any single dial are in all cases regions with equivalent final digit distributions.

A rigorous statistical analysis and interpretation of these data is limited by several conditions: (1) the occurrence of small error frequencies in many cells, in spite of (2) the fact that the tabulated frequencies represent the pooled data for all subjects, each of whom made a block of readings and read with a somewhat different error rate, and (3) the fact that individual scale readings vary in interpolation difficulty according to final digit. An analysis of variance of the data by subjects, even where the error frequencies might have permitted this, was passed over in favor of homogeneity tests on the pooled group data. Two such tests, both subject to certain limitations, were carried through: one based in Chi Square, wherever cell frequencies were adequate, the other based on the use of binomial probability paper, a recently available paper which permits a graphical analysis of frequency data via the angular transformation (see Mosteller and Tukey, 1949). Involved in the tests with binomial probability paper is the assumption that scale reading errors are binomially distributed. Although this is not strictly the case, previous analysis (Kappauf, 1947) has shown that the assumption is probably not too misleading in the treatment of dial reading data. To the extent that the assumption does not apply, that is, to the extent that error rates do vary by digit or subject or both, the binomial probability paper test becomes somewhat less sensitive than otherwise to deviant error frequencies within the pooled data. In the actual handling of the data on binomial paper, the Mosteller-Tukey modification of the angular transformation (for stabilizing variance at low and high percentages) was employed, and the range of plotted points was used as the test of homogeneity. The results of the application of these tests are shown in the tables by the symbols explained in the footnote to Table III.

Local Errors

Examination of the data in Section A of Tables II and III reveals no consistent error trend by sector or dial region. There are 22 dials in the two experiments where data by dial tenths or dial fifths were available, and among these there is but one for which the range of variation in number of errors by sector exceeds the 5% level, according to the graphical test. Over the two complete experiments there are 44 sets of records which may be considered by dial halves, right half vs. left half. For these, error differences of a size expected to occur one time in 20 by chance prove to be present three times in 44, according to each of the methods of testing used. Deviating differences thus occur no more often than would be anticipated if chance alone were responsible for the observed variations. Furthermore, they are divided in favoring right and left dial halves. Comparably negative results are obtained when error breakdowns are made by quadrants, right, left, upper and lower, similar to those used by Garner.

These tests of the data thus offer no evidence which would lead one to reject the hypothesis that local scale reading errors are equally likely in all parts of a dial viewed perpendicular to the reader's line of sight.

Errors of Five or More Units

For these errors, listed in Section B of the tables, the general picture is again negative. Trends consistent for all dials do not appear. The implication is that for dials in general, sector has no systematic effect on the frequency of large scale reading errors as here defined.

Analysis of the data, however, points to a further observation, that while dial region may have no uniform effect for the total class of dials and scales studied, it does influence errors made on the dials numbered by ten unit steps, the 50's and 100's dials. Among the 22 sets of dials in the two studies where records by dial tenths or dial fifths are tallied, there is one for which the range of variability of error frequencies by sector exceeds the 5% level and another for which it exceeds the 1% level according to the graphical test. Both are 50's dials. Under the same test, the 44 comparisons of dial halves are shown to include five differences at the 5% level and two others at the 1% level. Again, all of these are either 50's or 100's dials. Further, all involve an observed difference in the same direction, namely a difference in favor of greater reading accuracy on the left half of the dial than on the right. If one accepts the procedure of summing the data for all dials of the same scale range (e.g. all dials scaled to 50 units) in a given experiment, and this could be defended in view of the similar "large error" rates for such dials, the resulting summed errors for dial halves may be examined under Chi Square. The summed errors for the right and for the left halves of all the 100's dials in Experiment 3 are 29 and 14, which data yield a Chi Square value significant at the 5% level. The error sums for all 50's dials in Experiment 4 are 49 for the right half and 13 for the left, while the comparable sums for the 100's dials in this experiment are 44 and 10. These differences both yield values of Chi Square significant at the 1% level. The graphical tests agree in each instance. Both tests, however, show that the similarly computed sums for opposed halves of the 200's, 400's, or 600's dials are not significantly different. One is led to the conclusion, therefore, that sector influenced the performance of these subject groups differently when the dials being read were the 50's and 100's dials than when the dials being read had one of the longer scale ranges. Since the statistical tests call attention to the 50's and 100's dial data, it is appropriate to inspect the error records in more detail.

Experiment 1 (Kappauf, Smith, and Bray, 1947) dealt with 100's dials exclusively and, as mentioned above, gave preliminary evidence of differences in sector difficulty for these dials. A particularly high error rate was observed in the scale range from 0 to 9. When the present records are examined to determine what errors might be basic to the right half-left half differences for the 50's and 100's dials, it is found that the earlier results are confirmed not only with respect to the relatively high frequency of errors in the 0 to 9 range, but also with respect to the nature of the errors made there. In Experiment 1, 19 non-local errors were recorded for the scale interval from 0 to 9 and all of these involved an error of plus 10 units (±1). In Experiment 3 there were 13 errors of five units or more in this sector and again every one was an error of reading too large by 10 units (±1). In Experiment 4, this error accounted for 11 of the 15 misreadings on the 0 to 9 region on the 100's dials, and for 28 of 34 such errors on the 50's dials.

TABLE II

FREQUENCY OF ERRORS IN DIFFERENT DIAL SECTORS: EXPERIMENT 3

A: ERRORS OF ONE AND TWO UNITS

Dial	Diam.	Grad. by	Sector - by scale values									
			0-9	10-19	20-29	30-39	40-49	50-59	60-69	70-79	80-89	90-99
100	2.8"	10	8	8	9	10	4	8	9	4	7	9
		5	6	5	5	5	3	5	5	4	6	8
		2	2	2	1				1	1		1
		1	1	1		3		3	3	1	3	2
	1.4"	10	15	17	8	5	11	10	12	10	6	9
		5	7	7	7	4	9	5	3	3	7	6
		2	2	2	2	5	1	2	2			2
		1	4	5	6	6	4	4	5	6	7	3
		Totals:	45	47	38	38	32	37	40	29	36	40
			0-19	20-39	40-59	60-79	80-99	100-119	120-139	140-159	160-179	180-199
200	2.8"	10	26	22	20	23	19	22	24	23	18	19
		5	7	9	15	9	12	6	10	6	7	5b,c
		2	3	3	3	4	4	2	3	6	2	5
	1.4"	10	28	26	21	26	20	25	25	29	27	26
		5	19	18	21	22	18	11	21	14	19	14
		2	12	11	17	19	14	14	15	15	12	12
		Tøtals:	95	89	97	103	87	80	98	93	85	81

B: ERRORS OF FIVE OR MORE UNITS

Dial	Diam.	Grad. by	Sector - by scale values									
			0-9	10-19	20-29	30-39	40-49	50-59	60-69	70-79	80-89	90-99
100	2.8"	10	1	1							1	
		5		1	1	1				1		
		2			1			1				
		1		1	2			1				
	1.4"	10	2	1								
		5	5	1			1	5				
		2	2	1	1	1				1		1
		1	3	1	1			1		1	1	
		Totals:	13	7	6	2	1	8		3	2	1
			0-19	20-39	40-59	60-79	80-99	100-119	120-139	140-159	160-179	180-199
200	2.8"	10	1	5	3	3		1	1	1	2	
		5	5	1		5	3	4	1	3	4	
		2	1	2		2	1	4		3	3	3
	1.4"	10	3	2	3	1	2	1	2	2	3	2
		5	1	4	2	1	2	4	1	3	1	
		2		2	1	6		1	2		2	1
		Tøtals:	11	16	9	18	8	15	7	12	15	6

Note: Each cell of the table represents 80 readings.

TABLE III

FREQUENCY OF ERRORS IN DIFFERENT DIAL SECTORS: EXPERIMENT 4

Dial	Dial Diam.	Grad. by	A: Errors of one and two units — Sector - by scale values: 0-9	10-19	20-29	30-39	40-49		B: Errors of five or more units — Sector - by scale values: 0-9	10-19	20-29	30-39	40-49	
50	2.8 inch	10	26	18	22	14	20		6	3		1	1	*,b
		5	10	8	7	10	6		8	1	1	1	1	
		2	3		1	1			2	1	1			b
		1	1				1		2	1	2	1		
	1.4 inch	10	24	22	20	21	24		8	2	1		1**	,b
		5	5	5	7	5	1			2	1	1		
		2	4					b	3	2	1		1	
		1		6	1	1	2	*	5	1				bb

Dial	Dial Diam.	Grad. by	A: Right Half	A: Left Half	B: Right Half	B: Left Half
100	2.8 inch	10	76	68	4	1
		5	23	18	9	bb
		2	3	2	4	b
		1	3	2	5	b
	1.4 inch	10	102	89	4	1
		5	27	18	8	3
		2	9	4	3	
		1	6	13	7	5
200	2.8 inch	10	145	155	13	12
		5	52	40	13	21
		2	13	11	13	13
	1.4 inch	10	188	173	20	14
		5	96	80	19	15
		2	91	70	20	20
400	2.8 inch	10	242	246	23	29
		5	156	158	21	22
	1.4 inch	10	339	330	14	24
		5	277	292	18	22
600	2.8 inch	10	285	320 c	13	20
		5	240	275 b,c	33	31
	1.4 inch	10	364	375	17	20
		5	364	377	35	45

Note: Each cell for 50's dials represents 200 readings; for other dials, 600 readings.

Key to Symbols:

* 5% level) graphical test by dial fifths.
** 1% level)

b 5% level) graphical test by dial halves.
bb 1% level)

c 5% level, Chi Square test.

Inspection of the data shows that these misreadings typically involve pointer settings at the values 6, 7, 8, and 9. But it is further true that plus 10 units errors are more prevalent in the right dial quadrant which does not include the 0 to 9 region, than in the left quadrant. So the plus 10 units error is generally more common on the right halves of these dials than on the left.

A second but relatively unimportant error occurring dominantly on the right half of the 50's and 100's dials was the reversal error. The latter is defined simply as the error of reading the scale in the wrong direction from a numbered graduation (see Christensen, 1948). Reversal errors were very rare in the present experiments (see Table IV) but such as were observed were found principally on the 50's and 100's dials, where they occurred only on the right halves of the dials, and where two-thirds of them were again in the 0 to 9 region. That these reversal errors should be limited almost entirely to the 50's and 100's dials was unexpected, but their occurrence on the right halves of the dials is understandable in terms of the scale direction argument. It is on the right half of clockwise dials that the scale direction is opposed to that of conventional scales which proceed in increasing value upward.

TABLE IV

SUMMARY OF OCCURRENCES OF REVERSAL ERRORS

		Right half of dial	Left half of dial	Total readings each half
Experiment 3:	100's dials	3	-	3200
	200's "	-	-	2400
Experiment 4:	50's "	6	-	4000
	100's "	9	-	4800
	200's "	-	-	3600
	400's "	3	2	2400
	600's "	-	-	2400

To the foregoing remarks on 10 units errors on the 50's and 100's dials should be added one other observation, that these errors comprise a majority of all the large errors recorded on these dials. A comparably popular error does not appear in the data for the 200's, 400's, or 600's dials. It would appear, therefore, that the possibility of differences between the frequency of systematic reading errors made in various dial sectors is related to the diversity of errors committed on each dial. Such differences are less likely to occur for scales which are complex enough to involve systematic errors of many forms with potentially compensating sector distribution. When, however, a scale is simple enough that systematic errors are reduced to a few recognizable forms, sector differences in difficulty may appear as a function of factors which may concentrate these particular errors.

If these conclusions from pooled group data were based upon the performance of a single group of subjects, it would be proper to restrict their application to the future performance of those same subjects. But since the data from three different experiments with different subject groups are in basic agreement, it appears that the trends in question are characteristic of dial readers in general.

VI. DISCUSSION

Unclear from the present data is the specific reason for the demonstrated clustering of the large number of 10 units errors on the right half of the 50's and 100's dials. There was a confounding of several factors in the design and layout of the particular dials which were used here, so there are several possible explanations for the occurrence and location of the observed errors. It might be suggested (1) that systematic errors of all sorts are more frequent on the right half of dials bearing relatively simple scales; (2) that the plus 10 units error is characteristic of readings in low numerical regions on scales numbered by tens; (3) that this error is characteristic of readings on dials numbered by tens where there is no break at the zero point, that is, where continuous 360 degree markings of the scale might lead subjects to misconstrue the scale origin; (4) that this error is characteristic of dials numbered by tens which have their zero position at the top and on which the numbered divisions (10, 20, 30....) may be confused with conventional clock face numbers (1, 2, 3....); or (5) that this error is akin to the error made in reading combined coarse and fine scale instruments, and as such is influenced by scale direction or sector. At the moment, hypotheses (4) and (5) seem the most likely.

Hypothesis (5) implies that the reader makes wrong use of the scale numbers, taking the nearest scale number, rather than the next lower scale number, to which to add the value of the scale portion which the pointer has swept off between numbers. This would make the plus 10 units error on scales numbered by tens essentially the same as the plus 100 mil error in field artillery and the plus 1000 foot error in altimeter reading, where the reader clearly works from the nearest rather than the next lower value on the coarse scale (see Bray, 1948; Grether, 1947; Kappauf, 1949). The hypothesis that this form of error is more frequent on the right half of a clockwise dial than on the left is subject to a partial check in terms of data for the 200's dials and 600's dials. On these dials errors of plus one numbered division occurred occasionally, and while they split by dial halves in the expected direction, they are far from conclusive. Errors of plus one numbered division on right and left dial halves were in the ratio of 35:2 for the 50's dials and 43:7 for the 100's. They were in the ratio of 20:17 for the 200's dials and 7:0 for the 600's. It will be interesting to extend the data on this point, and one way of doing this will be through the examination of the records of other already published studies, as for example that of Christensen (1948).

That the number of large or systematic errors as defined in this paper was uncomfortably small for purposes of sector analysis is, of course, all too clear to the authors. To make an apology for low error frequencies, however, means only that all scale design research concerned with gross or systematic errors is the more difficult because it requires experiments based on extremely large numbers of readings. It is not necessarily an indication that the errors involved have rates of occurrence which are too low to be important in practical situations. By way of illustration, it may be pointed out that the total number of errors of plus 10 units made in reading pointer settings at 6, 7, 8, or 9 on all 50's and 100's dials was less than 20 for each of these settings. The error rates for these settings, however, were respectively 2.3%, 7.6%, 1.6%, and 4.4%. These figures, too, are probably to be interpreted as minimum values because they apply to a test situation where readers have become very familiar with the dial scales through the concentrated reading of dials with a given numerical range. Under conditions where the dial reading problem varies continually, higher error rates would be expected (Vernon, 1946).

VII. CONCLUSIONS

There is no evidence in the present data which would indicate that local scale reading errors are more likely in one dial region than in another.

There is evidence, on the other hand, that systematic errors are made with greater frequency on the right than on the left half of dials which cover a range of low scale values and are numbered by tens. The error most clearly identified with this right half-left half difference is the error of reporting a value which is too large by 10 units.

These data suggest that, all other things being equal, dials which range over low scale values and are numbered by tens should be so oriented that the scale region over which the most frequent and/or critical quantitative readings are made appears in the left dial half.

VIII. BIBLIOGRAPHICAL REFERENCES

Bray, C. W. Psychology and Military Proficiency: A history of the Applied Psychology Panel of the National Defense Research Committee. Princeton, N.J.: Princeton Univ. Press, 1948. pp xviii + 242. (Unclassified, English)

Christensen, J. M. The effect of the staircase scale on dial reading accuracy. USAF Memorandum Report No. MCREXD-694-1P, United States Air Force, Air Materiel Command, 1 October 1948. p 17. (Unclassified, English)

Garner, W. R. Some perceptual problems in the use of . . . . remote PPI. Systems Research Johns Hopkins Univ. September 15, 1946. p 34. (Restricted, English)

Grether, W. F. The effect of variations in indicator design upon speed and accuracy of altitude readings. USAF Memorandum Report No. TSEAA-694-14, United States Air Force, Air Materiel Command, 2 September 1947. p 19. (Unclassified, English)

Grether, W. F., and Connell, S. C. Psychological factors in check reading single instruments. USAF Memorandum Report No. MCREXD-694-17A, United States Air Force, Air Materiel Command, 20 September 1948. p 21. (Unclassified, English)

Kappauf, W. E. The use of the angular transformation in the statistical treatment of error frequencies. USAF Memorandum Report No. TSEAA-694-1J, 15 July 1947. (Unclassified, English)

KAppauf, W. E. Studies pertaining to the design of visual displays for aircraft instruments, computers, maps, charts, tables and graphs: A review of the literature. USAF Technical Report No. 5765, United States Air Force, Air Materiel Command, April 1949. p 100. (Unclassified, English)

Kappauf, W. E., and Smith, W. M. Design of instrument dials for maximum legibility: II. A preliminary experiment on dial size and graduation. USAF Memorandum Report No. MCREXD-694-1N, United States Air Force, Air Materiel Command, 12 July 1948. p. 16. (Unclassified, English)

Kappauf, W. E., Smith, W. M., and Bray, C. W. Design of instrument dials for maximum legibility: I. Development of methodology and some preliminary results. USAF Memorandum Report No. TSEAA-694-1L, United States Air Force, Air Materiel Command, 20 October 1947. p 42. (Unclassified, English)

Lebensohn, J. E, Scientific and practical considerations involved in the near-vision test with presentation of a practical and informative near-vision chart. Amer. J. Ophthal. 1936, 19. pp 110-117. (Unclassified, English)

Lewis, D., and Burke, C. J. The use and misuse of the chi-square test. Psychol. Bull. 1949, 46. pp 433-489. (Unclassified, English)

Mosteller, F., and Tukey, J. W. The uses and usefulness of binomial probability paper. J. Amer. Stat. Assn. 1949, 44. pp 174-212. (Unclassified, English)

Warrick, M. J., and Grether, W. F. The effect of pointer alignment on check reading of engine instrument panels. USAF Memorandum Report No. MCREXD-694-17, United States Air Force, Air Materiel Command, 4 June 1948. p 33. (Unclassified, English)

Vernon, M. D. Scale and dial reading. Unit in Applied Psychology, the Psychological Laboratory, Cambridge, A.P.U. 49. June 1946. (Unclassified, English)

ADDENDUM

Since the preparation of this manuscript, the following Technical Reports dealing with errors in radar scope reading have come to the attention of the writers:

Horton, G. P. An analysis of errors made in a schematic PPI display. USAF Technical Report No. 5960, United States Air Force, Air Materiel Command, October 1949a. (Unclassified, English)

Horton, G. P. Accuracy of reading target location and size of schematic PPI display. USAF Technical Report No. 5961, United States Air Force, Air Materiel Command, October 1949b. (Unclassified, English)

In one breakdown in the first of these articles, Horton considers systematic errors in readings of target bearing (defined as errors greater than five degrees which involve an incorrect tens or hundreds digit) as a function of scope sector. The bearing scale used on his simulated scope display was numbered every 10 degrees from 0 to 360 and was graduated every two, thus resembling the bearing scales used on many PPI's. The display was presumably viewed perpendicular to the reader's line of sight. Error data are presented by 20 degree sectors and represent a total of 500 readings per sector. The total number of readings over all sectors approximates the number reported for all 400's and 600's dials combined in Experiment 4 above.

Highly significant in Horton's data is the increase in the frequency of systematic errors with sector. By 90 degree quadrants, starting from zero degrees, these errors increased in frequency from 3.2% to 4.6%, to 7.5&, to 8.8%. Thus they were more than twice as frequent on the left half of the scope than on the right. Although the data from Experiment 4 for 400's and 600's dials show a trend toward more frequent systematic errors (some 25% more) on the left dial half than on the right, this difference is not statistically significant in the case of any single dial. Inasmuch as the data for Experiments 3 and 4 combined have been interpreted to mean that sector effects in regard to systematic errors of scale reading will not be present unless some particular error (or group of errors) dominates the error pattern, the question to be raised is whether Horton's reading conditions were such as to give rise to some specific form of error which would increase in likelihood with indicated bearing or azimuth. One can only make conjectures at this point, but since these hunches may be checked through a re-analysis of certain other data which Horton has reported upon (second reference above), they should be worth noting here.

The data which Horton analyzed by sector were for target readings made during five second scope exposures. Subjects were instructed to report both target range and bearing, giving range first. These tests thus had the flavor of a tachistoscopic experiment and the observed reading errors were more frequent than if reading time had not been restricted. The reading conditions were further characterized by the general unfamiliarity of the zero to 360 degree bearing scale and by the fact that the successive 10 degree numbers were uniform and unaccented. It is therefore reasonably certain that Horton's scope readers were obliged to read the scale numbers in taking data on almost every target. They could not rely very much on known values of familiar or outstanding reference points as they do in reading familiar scales like a clock face. How likely is it, then, that Horton's trend of errors with sector was the result of differences in confusability or reading difficulty of the scale numbers under the imposed conditions of limited reading time? The probability seems quite good. On the right half of the scope the numbers are more adequately spaced than on the left. And within the zero to 90 degree quadrant the tens digits, which are critically important to successful reading, are not lost in the middle of the numbers but define their external contours in a way which should make the numbers more readily recognized.

A test of this "number reading hypothesis" to account for Horton's data can be made through an examination, by sector, of the systematic errors made when subjects are given unlimited time to make their readings. Such data have been gathered by Horton in the course of his extended studies, and it is to be hoped that he will be checking them also for sector effects.

In the meantime, there is some evidence in favor of at least a portion of the present hypothesis. Familiar anchor points on the number scale should, according to the above notions, reduce the occurrence of systematic errors within sectors containing them. Two such points are more familiar than the rest of the bearing number scale, namely the 90 and 270 degree points. And as Horton has himself noted, the sectors containing these points do have fewer errors than immediately adjacent sectors.

If the present hunches about the importance of scale number reading are borne out, it is clear that they should be taken into account in the design of bearing scales in general. Lacking the experience-structured advantages of more familiar scales, or the simplicity of scales with fewer numbers, bearing scales should be structured as adequately as possible through the use of more conspicuously accented numbers and graduation marks.

It might also be suggested that, if the remarks in this addendum are borne out, the principal purpose of the present report will have been indicated -- namely, to call attention to the specific nature of sector effects and to prompt investigations into their particular origins on those scales for which they are reported.

TITLE: Design of Instrument Dials for Maximum Legibility - Part 4. Dial Graduation, Scale Range and Dial Size as Factors Affecting the Speed and Accuracy of Scale Reading - and Appendix (AF Technical Report)	ATI-Part of 78 314
AUTHOR(S): Kappauf, William E.; Smith, William M.	DIVISION (None)
ORIG. AGENCY: Princeton University, N.J.	ORIG. AGENCY NO. (None)
PUBLISHED BY: AMC, Wright-Patterson Air Force Base, Dayton, O.	PUBLISHING AGENCY NO. AF-TR-5914, Part 4

DATE	U.S. CLASS.	COUNTRY	LANGUAGE	PAGES	ILLUSTRATIONS
Feb' 50	Unclass.	U.S.	English	21	tables, graphs

ABSTRACT:

The extent to which instrument-dial reading errors and reading times are influenced by the spacing of graduation marks, graduation-mark values, scale range, and dial size was investigated. The dials employed were scaled from 0 to 50, 0 to 100, 0 to 200, O to 400, and O to 600 units, and included graduations by tens, fives, twos, and units. Dial sizes were 2.8 in. and 1.4 in. The precision of scale reading varied with the distance on the scale allocated to each scale unit. Reading precision improved as this distance increases to about 0.05 in., a value which varied only slightly with gradution scheme but which appeared to vary with graduation mark thickness. For more expanded scales, reading precision remained reasonably uniform at a level which depends on graduation scheme.

DISTRIBUTION: Copies of this report obtainable from CADO.

DIVISION: Aviation Medicine (19)
SECTION: General (0)

SUBJECT HEADINGS: Instrument dials - Legibility

USAF Contr. No. W33-038-ac-14480